

METHODS FOR DETERMINING POLYPEPTIDE STRUCTURE, FUNCTION
OR PHARMACOPHORE FROM COMPARISON OF POLYPEPTIDE SEQUENCES

ABSTRACT OF THE DISCLOSURE

The invention provides a method for separating
5 two or more subsets of polypeptides within a set of
polypeptides. The method includes the steps of: (a)
determining a sequence comparison signature for each
amino acid sequence in a set of amino acid sequences,
wherein the sequence comparison signature includes
10 pairwise comparison scores for the amino acid sequence
compared to each of the other amino acid sequences in the
set; (b) constructing a distance arrangement including
the sequence comparison signatures related according to
the distance between each of the sequence comparison
15 signatures; and (c) identifying a first and second
cluster of sequence comparison signatures in the distance
arrangement, wherein the first cluster includes sequence
comparison signatures for polypeptides having a similar
protein fold or biological function, the protein fold or
20 function being different compared to a protein fold or
function of polypeptides having sequence comparison
signatures in the second cluster.